

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 15:43:18 : Search time 63.5263 Seconds
(without alignments)
2848.398 Million cell updates/sec

Title: US-09-974-973A-19

Perfect score: 5788

Sequence: 1 MSTHTSSTLPFRKILVANR.....RVVYPATKVGGLIIVVS 1140

Scoring table:
BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.GeneSeq_19Jun03.*

```

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5788	100.0	1140	22	AAG90511
2	5788	100.0	1140	22	AAB67129
3	5788	100.0	1140	23	AAE25601
4	5788	100.0	1140	23	AAU98053
5	5784	99.9	1141	21	AAB01436
6	5782	99.9	1140	20	AAW35971
7	5780	99.9	1140	22	AAW35971
8	5759	99.5	1140	23	AAU98052
9	5759	99.5	1157	23	AAU98050

10	5302.5	91.6	1139	22	AAB83180
11	2621	45.3	532	22	AAB79302
12	2524	43.6	1148	22	AAU00511
13	2493	43.1	1146	23	ABB47612
14	2488	43.0	1147	22	AAU33972
15	2472.5	42.7	1142	22	AAU35213
16	2464	42.6	1151	22	AAU36877
17	2464	42.6	1154	22	ABP35883
18	2441	42.2	1181	22	ABB58211
19	2441	42.2	1181	22	ABB66604
20	2441	42.2	1181	22	ABB66605
21	2433.5	42.0	1196	22	ABB67309
22	2405	41.6	1136	23	ABB53980
23	2357	40.7	461	22	AAB79303
24	2299	39.7	1073	22	AAU36768
25	1546.5	26.7	320	22	AAB79300
26	1546.5	26.7	320	22	AAB79301
27	1375	23.8	272	22	AAB79298
28	1375	23.8	272	22	AAB79299
29	1036.5	17.9	447	17	AAW05207
30	1036.5	17.9	447	19	AAW70402
31	1036.5	17.9	448	15	AAR51080
32	1004.5	17.4	453	15	AAR51083
33	1004.5	17.4	453	17	AAW05208
34	1004.5	17.4	453	19	AAW70403
35	978	16.9	425	21	AAB11663
36	976.5	16.9	471	22	AAU33719
37	957	16.5	449	22	AAU38292
38	953.5	16.5	456	23	ABP28017
39	951.5	16.4	455	23	ABB54091
40	946	16.3	448	22	AAU35567
41	942.5	16.3	455	22	AAU38024
42	942.5	16.3	455	24	ABU00779
43	941.5	16.3	455	22	AAU37771
44	939.5	16.2	455	22	AAW01054
45	938	16.2	449	16	AAR66742

ALIGNMENTS

```

RESULT 1
AAG90511
ID: AAG90511 strand: Protein; 1140 AA.
XX
AC AAG90511;
XX
26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 4265.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99AP-0377484.
XX 07-APR-2000; 2000UP-0159162.
XX 03-AUG-2000; 2000UP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI: 2001-376931/40.
XX
DR N-PSDB; AAH65730.

```

Corynebacterium th
Corynebacterium gl
Bacillus subtilis
Listeria monocytog
Staphylococcus aur
Enterococcus faeca
S. epidermidis ope
Staphylococcus epi
Drosophila melanog
Drosophila melanog
Drosophila melanog
Lactococcus lactis
Corynebacterium gl
Staphylococcus aur
Corynebacterium gl
Corynebacterium gl
Corynebacterium gl
Anabaena biotin bi
Anabaena biotin ca
biog gene encoding
Synecococcus biot
A. vitis hyperiens
Pseudomonas aerugi
Salmonella typhi c
Streptococcus poly
Lactococcus lactis
Haemophilus influe
Streptococcus pneu
S. pneumoniae type
Streptococcus pneu
CPE 57 protein seq
Biotin-carboxylase


```

XX Sequence 1140 AA:
SQ Query Match 100.0%; Score 5788; DB 22; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPFAFKILVANRGEIAVRAFAALETGAATVAIYPREDGSHRSPASEAVR 60
DB 1 MSTHTSSTLPFAFKILVANRGEIAVRAFAALETGAATVAIYPREDGSHRSPASEAVR 60

QY 61 IGTGSPVYKAYLDIDELITGAKKVKADAIYPGYGFLENNQOLARECAENGITTFGPTPEV 120
DB 61 IGTGSPVYKAYLDIDELITGAKKVKADAIYPGYGFLENNQOLARECAENGITTFGPTPEV 120

QY 121 LDLTGDKSRVAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFYKAVAGGGGRGMRF 180
DB 121 LDLTGDKSRVAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFYKAVAGGGGRGMRF 180

QY 181 VASPDRLKLTATASREAEAFDGAAYVERAVINPHIEVOILGDHTGEVHLXYERDCS 240
DB 181 VASPDRLKLTATASREAEAFDGAAYVERAVINPHIEVOILGDHTGEVHLXYERDCS 240

QY 241 LQRRHOKVVEIAPAOHLDPRLRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEMN 300
DB 241 LQRRHOKVVEIAPAOHLDPRLRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEMN 300

QY 301 PRIQVEHTVEEYEVNDLVKAQKRLAAGATLKEIGLTDQKIKTHGAALQCRITTEDPNNG 360
DB 301 PRIQVEHTVEEYEVNDLVKAQKRLAAGATLKEIGLTDQKIKTHGAALQCRITTEDPNNG 360

QY 361 FRPDGTGTTAYRSPGGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAARAQAL 420
DB 361 FRPDGTGTTAYRSPGGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAARAQAL 420

QY 421 AEFYVSGVATNIGFLRALNREDEFTSKRIATGFIADPHILOAPPADDEGRILYDLADY 480
DB 421 AEFYVSGVATNIGFLRALNREDEFTSKRIATGFIADPHILOAPPADDEGRILYDLADY 480

QY 481 TVKKPHGVKPKDVAAPLDKLPNTKOLPLPGSRDLKOLCPAARLREODALAVTDYT 540
DB 481 TVKKPHGVKPKDVAAPLDKLPNTKOLPLPGSRDLKOLCPAARLREODALAVTDYT 540

QY 541 FRDAHOSILATRVRSFALPAEAVAKLPELISVEAMGGATDVAMRFLFEDPMWDLDE 600
DB 541 FRDAHOSILATRVRSFALPAEAVAKLPELISVEAMGGATDVAMRFLFEDPMWDLDE 600

QY 601 LREAMPNVNIQMLLRGNITVGYTPYDPSVCRAFEVKEAASSGVDFIRIFDALNDVSOQMPA 660
DB 601 LREAMPNVNIQMLLRGNITVGYTPYDPSVCRAFEVKEAASSGVDFIRIFDALNDVSOQMPA 660

QY 661 IDAVLETNNAVAEVAAYSGDLSDPNEKLYTLDYYLKAEBEIVKSGAHILAIKDMACLLR 720
DB 661 IDAVLETNNAVAEVAAYSGDLSDPNEKLYTLDYYLKAEBEIVKSGAHILAIKDMACLLR 720

QY 721 PAAVTKLVTLAREEFDLPVHVHTHDAGGOLATYFAAQAQGAADVADASAPLSGTTSOPS 780
DB 721 PAAVTKLVTLAREEFDLPVHVHTHDAGGOLATYFAAQAQGAADVADASAPLSGTTSOPS 780

QY 781 LSAIVAAFAHTRBDTGLSEAVSDLEPYWAVNRGLYLPFSGTPGPGRYRHEIPSGOL 840
DB 781 LSAIVAAFAHTRBDTGLSEAVSDLEPYWAVNRGLYLPFSGTPGPGRYRHEIPSGOL 840

QY 841 SNLRQAOTALGLADREFLIDNVAANVEMLGPRPKYTPSSKVVGDLLHLVGACVPADP 900
DB 841 SNLRQAOTALGLADREFLIDNVAANVEMLGPRPKYTPSSKVVGDLLHLVGACVPADP 900

QY 901 AADPOKIDIPDSVIAFLRGELGNPGGWPEPLRTALRSEKKAPLTEVPEEEOALHDA 960
DB 901 AADPOKIDIPDSVIAFLRGELGNPGGWPEPLRTALRSEKKAPLTEVPEEEOALHDA 960

QY 961 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDREFFYGLVEGRETLIRLPDVRTP 1020
DB 961 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDREFFYGLVEGRETLIRLPDVRTP 1020

```

```

DB 961 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDREFFYGLVEGRETLIRLPDVRTP 1020
QY 1021 LTVRLDAISRPDDKGMNRVNAVNGOIRPMKRVRSVESYTATPERKADSSKNGHVAAPFA 1080
DB 1021 LTVRLDAISRPDDKGMNRVNAVNGOIRPMKRVRSVESYTATPERKADSSKNGHVAAPFA 1080
QY 1081 GVYTVTAEBDEVACGAVAIIEAMKMEATITASVDKIDRVVVPATKYEGGDLIYVVS 1140
DB 1081 GVYTVTAEBDEVACGAVAIIEAMKMEATITASVDKIDRVVVPATKYEGGDLIYVVS 1140

RESULT 3
AAE25601
ID AAE25601 standard; Protein; 1140 AA.
XX
AC AAE25601;
XX
DT 04-NOV-2002 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase protein.
XX
KW Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;
XX oxaloacetate; growth; enzyme.
XX
OS Corynebacterium glutamicum.
XX
FH Key location/qualifiers
FT Misc-difference 1 /note= "Encoded by GTG"
XX
PN US6403351-B1.
XX
PD 11-JUN-2002.
XX
PE 03-OCT-2000; 2000US-0677575.
XX
PR 23-DEC-1998; 98US-0220081.
XX
PA (ARCH ) ARCHER-DANTELS MIDLAND CO.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR MPI: 2002-536037/57.
XX
DR N-PSDB: AAD42046.
XX
PT Novel pyruvate carboxylase polypeptide, useful for replenishing
PT oxaloacetate consumed for biosynthesis during growth, or lysine and
PT glutamic acid production in industrial fermentation -
PS Claim 1; Column 29-36; 28pp; English.
XX
CC The present invention relates to novel pyruvate carboxylase proteins and
CC polynucleotides encoding such proteins. Sequences of the invention are
CC important anaplerotic enzymes for replenishing oxaloacetate consumed for
CC biosynthesis during growth, or lysine and glutamic acid production in
CC industrial fermentation. The present sequence is C. glutamicum pyruvate
CC carboxylase protein.
XX
SQ Sequence 1140 AA:
Query Match 100.0%; Score 5788; DB 23; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPFAFKILVANRGEIAVRAFAALETGAATVAIYPREDGSHRSPASEAVR 60
DB 1 MSTHTSSTLPFAFKILVANRGEIAVRAFAALETGAATVAIYPREDGSHRSPASEAVR 60

QY 61 IGTGSPVYKAYLDIDELITGAKKVKADAIYPGYGFLENNQOLARECAENGITTFGPTPEV 120
DB 61 IGTGSPVYKAYLDIDELITGAKKVKADAIYPGYGFLENNQOLARECAENGITTFGPTPEV 120

QY 121 LDLTGDKSRVAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFYKAVAGGGGRGMRF 180
DB 121 LDLTGDKSRVAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFYKAVAGGGGRGMRF 180

```

Db 121 LDLTGDKSRAYTAAKRAGLPVLAESTPSKNIDELVKSAGEQPIEFVKAAGGGRMFR 180
QY 181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIIEVOILGHTGEVHLHYERDCS 240
Db 181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIIEVOILGHTGEVHLHYERDCS 240
QY 241 LORRHOKVEIAPAOHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEM 300
Db 241 LORRHOKVEIAPAOHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEM 300
QY 301 PRIQVHTVTEETVEVDLYKAOMRLAAGATLKEGLTQDKIKTHGAALOCRTTEDPNNG 360
Db 301 PRIQVHTVTEETVEVDLYKAOMRLAAGATLKEGLTQDKIKTHGAALOCRTTEDPNNG 360
QY 361 FRPDGTITAYRSPGAGYRLDGAOLGGEITAHFDSMLVKMTGCRSDDEETANARORAL 420
Db 361 FRPDGTITAYRSPGAGYRLDGAOLGGEITAHFDSMLVKMTGCRSDDEETANARORAL 420
QY 421 AEFTVSGVATNIGFLRALLREDEFTSKRIATGFIADHPHLLQAPPADDEGRITLDYLA 480
Db 421 AEFTVSGVATNIGFLRALLREDEFTSKRIATGFIADHPHLLQAPPADDEGRITLDYLA 480
QY 481 TVNKPGRKDVAPKDAIDKLPNKDLPRLGRSGHRLKQIGRAAFADLKEODALAVTPTT 540
Db 481 TVNKPGRKDVAPKDAIDKLPNKDLPRLGRSGHRLKQIGRAAFADLKEODALAVTPTT 540
QY 541 FRDAGSLATRYRSFALRPAEAVALKLPPELLSVAMGAGATYDVMARLFEDEPMRLDE 600
Db 541 FRDAGSLATRYRSFALRPAEAVALKLPPELLSVAMGAGATYDVMARLFEDEPMRLDE 600
QY 601 LREAMPNNIOMLRGNVTGYTPYDVSRAAFVKEAASSGVDIFRIFDALNDVSGMRPA 660
Db 601 LREAMPNNIOMLRGNVTGYTPYDVSRAAFVKEAASSGVDIFRIFDALNDVSGMRPA 660
QY 661 IDAVLETNVAEVAAYSGDLSDPNEKIYTDYIKMAEBEYKSAHILAIKDMAGLIR 720
Db 661 IDAVLETNVAEVAAYSGDLSDPNEKIYTDYIKMAEBEYKSAHILAIKDMAGLIR 720
QY 721 PAAVTKLVTLRREFDLPVHVTHTDYGQGLATYFAAQAQADAVDAGASAPISGTTSPS 780
Db 721 PAAVTKLVTLRREFDLPVHVTHTDYGQGLATYFAAQAQADAVDAGASAPISGTTSPS 780
QY 781 LSAIYAAFAHTRDGLSLAVALSDELPYEAARGLYLPRESGTPGTGVRHHEIPGGOL 840
Db 781 LSAIYAAFAHTRDGLSLAVALSDELPYEAARGLYLPRESGTPGTGVRHHEIPGGOL 840
QY 841 SNLRQAATLGLADREFELIEDNYAANEMLGRPTKVTPTSSKVVGDALHLVAGADPADF 900
Db 841 SNLRQAATLGLADREFELIEDNYAANEMLGRPTKVTPTSSKVVGDALHLVAGADPADF 900
QY 901 AADPOKYDIPDSVIAFLRGELGNPPGMEPLRTRALBGRSGKAPLITEVPEEQAHDLDA 960
Db 901 AADPOKYDIPDSVIAFLRGELGNPPGMEPLRTRALBGRSGKAPLITEVPEEQAHDLDA 960
QY 961 DDKSKERRNLNLPPKPTPEEFLEHRRRGNTSALDDREFYGLVGEGETLIRLDVPTP 1020
Db 961 DDKSKERRNLNLPPKPTPEEFLEHRRRGNTSALDDREFYGLVGEGETLIRLDVPTP 1020
QY 1021 LTVRLDAISEPDDKGRNVAVVNGOIRPMRYRDSVESVTATAKADSSNKGHAAPPA 1080
Db 1021 LTVRLDAISEPDDKGRNVAVVNGOIRPMRYRDSVESVTATAKADSSNKGHAAPPA 1080
QY 1081 GVVTYVVAEGDEKAGDAVAIIEAKKMEATITASVDGKIDRVVVAATKVEGGLIYVVS 1140
Db 1081 GVVTYVVAEGDEKAGDAVAIIEAKKMEATITASVDGKIDRVVVAATKVEGGLIYVVS 1140

RESULT 4
AAU98053
ID AAU98053 standard; Protein; 1140 AA.
XX
AC AAU98053;

XX 27-AUG-2002 (first entry)
DT Corynebacterium wild-type feedback-resistant pyruvate carboxylase enzyme.
XX
DE Feedback-resistant; pyruvate carboxylase; enzyme;
XX aspartic acid feedback inhibition resistant.
KW
XX Corynebacterium glutamicum.
OS
PN WO200231158-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001MO-US31893.
XX
PR 13-OCT-2000; 2000US-239913P.
XX
PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Hanke PD;
XX
PT Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT polypeptide, useful for producing amino acids e.g. L-lysine,
PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PS L-isoleucine
PS Disclosure; Fig 2; 42pp; English.
XX
CC The present invention relates to a new mutated, feedback-resistant
CC pyruvate carboxylase enzyme. The invention is useful for producing an
CC amino acid (e.g. L-lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC by culturing a host cell in a suitable media and separating the amino
CC acid from the medium. The vector of the invention is useful for
CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC a selectable marker gene through homologous recombination to form a first
CC recombinant strain, and replacing the selectable marker gene in the
CC first recombinant strain, with feedback resistant pyruvate carboxylase
CC gene through homologous recombination to form a second recombinant
CC strain, where the homologous recombination in the above steps, occurs
CC between the host cell and the vector. The feedback-resistant pyruvate
CC carboxylase enzyme is resistant to feedback inhibition from aspartic
CC acid. The present amino acid sequence represents the wild-type
CC feedback-resistant pyruvate carboxylase enzyme of the invention.
XX
SQ Sequence 1140 AA;
Query Match 100.0%; Score 5788; DB 23; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTITSSTLPAFFKILVANRGEIARAFRAALETGAATVAIYPRDRGSFHRSEASEAVR 60
Db 1 MSTITSSTLPAFFKILVANRGEIARAFRAALETGAATVAIYPRDRGSFHRSEASEAVR 60
QY 61 IGTESSPVKAYLIDDEITGAARKKVAADAIYGYGLSLNQAOLARCAENGITFTGPTPEV 120
Db 61 IGTESSPVKAYLIDDEITGAARKKVAADAIYGYGLSLNQAOLARCAENGITFTGPTPEV 120
QY 121 LDITGDKSRAYTAAKKAGLPVLAESTPSKNIDELVKSAGEQPIEFVKAAGGGRMFR 180
Db 121 LDITGDKSRAYTAAKKAGLPVLAESTPSKNIDELVKSAGEQPIEFVKAAGGGRMFR 180
QY 181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIIEVOILGHTGEVHLHYERDCS 240
Db 181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIIEVOILGHTGEVHLHYERDCS 240
QY 241 LORRHOKVEIAPAOHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEM 300
Db 241 LORRHOKVEIAPAOHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFTEM 300

Db 241 LQRRHOKVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 300
 Qy 301 PRIOVEHTYEETEVLDVKAQMRILAAGATLKLGLTQOKIKTHGAALQCRITTEDPNNG 360
 Db 301 PRIOVEHTYEETEVLDVKAQMRILAAGATLKLGLTQOKIKTHGAALQCRITTEDPNNG 360
 Qy 361 FRPDGTITAYRSPGAGVRLDGAOALGGEITAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Db 361 FRPDGTITAYRSPGAGVRLDGAOALGGEITAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Qy 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILLDYADV 480
 Db 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILLDYADV 480
 Qy 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDIT 540
 Db 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDIT 540
 Qy 541 FRDAHOSLITAYRSEFALKPAAEAVAKLTPPELLSVEAMGATYDVAMRFLFEDPMRLDE 600
 Db 541 FRDAHOSLITAYRSEFALKPAAEAVAKLTPPELLSVEAMGATYDVAMRFLFEDPMRLDE 600
 Qy 601 LREAMPNVNIQMLLRGNITGYTPYDSCRAEYKKAASSGVDFIRILDALNDVSOQMPA 660
 Db 601 LREAMPNVNIQMLLRGNITGYTPYDSCRAEYKKAASSGVDFIRILDALNDVSOQMPA 660
 Qy 661 IDAVLENTNVAEVAAMAYSGDSLDPNKEKLTLDLYLKMAGEIYKSGHIIAIDMAGLLR 720
 Db 661 IDAVLENTNVAEVAAMAYSGDSLDPNKEKLTLDLYLKMAGEIYKSGHIIAIDMAGLLR 720
 Qy 721 PAATKLTVALRREFDLPVVHTHDTAGQLATYFAAAQAGADAVDASAPLSGTTSQPS 780
 Db 721 PAATKLTVALRREFDLPVVHTHDTAGQLATYFAAAQAGADAVDASAPLSGTTSQPS 780
 Qy 781 LSAIVAAFAHTRDGTGLSEAVSDLEPYEAVRGVLYPESGTPGPTGRVYRHEIPGGQL 840
 Db 781 LSAIVAAFAHTRDGTGLSEAVSDLEPYEAVRGVLYPESGTPGPTGRVYRHEIPGGQL 840
 Qy 841 SNIRAOATLGLDREFLEDNYAAVNMELGPRTKVPSKVVGDALHLVAGVDPADF 900
 Db 841 SNIRAOATLGLDREFLEDNYAAVNMELGPRTKVPSKVVGDALHLVAGVDPADF 900
 Qy 901 AADPOKYDIPDSYIAFLRGEIAGNPGMPERPLTRALEGRSEKAPLTEVEPEEOAHIDA 960
 Db 901 AADPOKYDIPDSYIAFLRGEIAGNPGMPERPLTRALEGRSEKAPLTEVEPEEOAHIDA 960
 Qy 961 DSKERRNSINRLLPKPTPEELHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVTRP 1020
 Db 961 DSKERRNSINRLLPKPTPEELHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVTRP 1020
 Qy 1021 LTVRLDAISRPDDKGMNVVAVANGQIRPMRVDRDSVESYTAATEKADSSNKGIVAAPFA 1080
 Db 1021 LTVRLDAISRPDDKGMNVVAVANGQIRPMRVDRDSVESYTAATEKADSSNKGIVAAPFA 1080
 Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDKIDRVVVPATKTEGGDLIYVVS 1140
 Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDKIDRVVVPATKTEGGDLIYVVS 1140

RESULT 5
 AAB01436
 ID AAB01436 standard; Protein: 1141 AA.
 AC AAB01436;
 XX 20-OCT-2000 (first entry)
 DE Pyruvate carboxylase of C. glutamicum.
 XX
 KW Pyruvate carboxylase; expression: amino acid biosynthesis; lysine;
 XX glutamic acid; oxaloacetate; fermentation; biosynthesis.
 OS Corynebacterium glutamicum.

XX
 PN W0200039305-A1.
 PD 06-JUL-2000.
 XX
 PF 23-DEC-1998; 98WO-US27301.
 PR 23-DEC-1998; 98WO-US27301.
 PA (SINS/) SINSKEY A J.
 PA (LESS/) LESSARD P A.
 PA (WILL/) WILLIS L B.
 PI Sinskey AJ, Lessard PA, Willis LB;
 DR WPI: 2000-465746/40.
 DR N-PSDB: AAA47533.
 XX
 PT Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
 PT carboxylase useful for industrial fermentation processes comprises a
 PT specific nucleotide sequence
 PS
 PS Claim 3; Fig 1; 51pp; English.
 CC The pyruvate carboxylase of Corynebacterium glutamicum can be used
 CC for producing amino acids, preferably lysine and glutamic acid in
 CC industrial fermentations and for replenishing oxaloacetate consumed
 CC for biosynthesis during growth. By incorporating the pyruvate
 CC carboxylase gene in expression vectors levels of expression can be
 CC 2 - 20 fold higher than in Corynebacterium glutamicum.
 XX
 SQ Sequence 1141 AA:
 Qy 1 MSPTSTLPAFKKILVANGELVRAFALEGAATVATYPRDGSFHRSPASAVR 60
 Db 2 VSTHTSSLPFAFKKILVANGELVRAFALEGAATVATYPRDGSFHRSPASAVR 61
 Qy 61 IGTEGSPVAKYLDIDELIIGAAKKVKAADATPYGGLSENAQARECAENGITFGPTPEV 120
 Db 62 IGTEGSPVAKYLDIDELIIGAAKKVKAADATPYGGLSENAQARECAENGITFGPTPEV 121
 Qy 121 LDLTGDKSRVATAAKKAGLPYLAESTPSKINDELIVKSAEGQYPIFKAAVAGGGRGMRP 180
 Db 122 LDLTGDKSRVATAAKKAGLPYLAESTPSKINDELIVKSAEGQYPIFKAAVAGGGRGMRP 181
 Qy 181 VASPEDELRLKLTAEASREAEAFDGAAYVERAVINPQHIIEYQILGDHTGEVYHLHERDCS 240
 Db 182 VASPEDELRLKLTAEASREAEAFDGAAYVERAVINPQHIIEYQILGDHTGEVYHLHERDCS 241
 Qy 241 LQRRHOKVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 300
 Db 242 LQRRHOKVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 301
 Qy 301 PRIOVEHTYEETEVLDVKAQMRILAAGATLKLGLTQOKIKTHGAALQCRITTEDPNNG 360
 Db 302 PRIOVEHTYEETEVLDVKAQMRILAAGATLKLGLTQOKIKTHGAALQCRITTEDPNNG 361
 Qy 361 FRPDGTITAYRSPGAGVRLDGAOALGGEITAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Db 362 FRPDGTITAYRSPGAGVRLDGAOALGGEITAHFDSMLVKMTCRGSDFETAVARAORAL 421
 Qy 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILLDYADV 480
 Db 422 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILLDYADV 481
 Qy 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDIT 540
 Db 482 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDIT 541

QY 541 FRDAHQSLATRVRSFALKPAAEAVALKLPBELLSEVAMGATYDVAMRFEDPWRDLDE 600
|||||
DB 542 FRDAHQSLATRVRSFALKPAAEAVALKLPBELLSEVAMGATYDVAMRFEDPWRDLDE 601
DB 601 LREAMPNVNIOMLLGRNTVGYTPYDSCRAFYKAEASSGVDFIRIFDALNDVSQMRPA 660
|||||
QY 661 IDAVLENTTAAVAEAMAYSGDLSDPNEKLYTDYLLKMAEETVKSAGHILAIKDMAGLLR 720
|||||
DB 662 IDAVLENTTAAVAEAMAYSGDLSDPNEKLYTDYLLKMAEETVKSAGHILAIKDMAGLLR 721
QY 721 PAAVTKLVLTALRREDDLPVHVHTHTDTAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780
|||||
DB 722 PAAVTKLVLTALRREDDLPVHVHTHTDTAGGOLATYFAAAGADAVDASAPLSGTTSQPS 781
QY 781 LSAIYAAFAHTRRDGLSLAEAVSDLEPYWEAVRGILYLPESGTPGPTGVRVHHEIPGGOL 840
|||||
DB 782 LSAIYAAFAHTRRDGLSLAEAVSDLEPYWEAVRGILYLPESGTPGPTGVRVHHEIPGGOL 841
QY 841 SNLRAQATATGLADRFELIEDNYAAVNEMIGRPTKVTPTSSKVVGDALHLVAGVDPADP 900
|||||
DB 842 SNLRAQATATGLADRFELIEDNYAAVNEMIGRPTKVTPTSSKVVGDALHLVAGVDPADP 901
QY 901 AADPOKYIDIPDSVIAFLKRGELGNPPGMPPEPLRTALBGRSEKAPLFEVPEEQAHLLDA 960
|||||
DB 902 AADPOKYIDIPDSVIAFLKRGELGNPPGMPPEPLRTALBGRSEKAPLFEVPEEQAHLLDA 961
QY 961 DDSKRRRSINLRLPPKPTPEEFLRHRRRRGNTSALDDREFFGVLGEGRETLRLRDPVRRP 1020
|||||
DB 962 DDSKRRRSINLRLPPKPTPEEFLRHRRRRGNTSALDDREFFGVLGEGRETLRLRDPVRRP 1021
QY 1021 LLYRLDAISEPDDCKMRNVANVNGQIRPMRVRSVS SVTATAEKADSSNKGHVAAPPA 1080
|||||
DB 1022 LLYRLDAISEPDDCKMRNVANVNGQIRPMRVRSVS SVTATAEKADSSNKGHVAAPPA 1081
QY 1081 GVVTYTVABEGDEVKAGDAVAITTEAMKMEATTITASVDGKITDRVYVAAATKVEGGDLTVVVS 1140
|||||
DB 1082 GVVTYTVABEGDEVKAGDAVAITTEAMKMEATTITASVDGKITDRVYVAAATKVEGGDLTVVVS 1141

RESULT 6
AAM93971
ID AAM93971 standard; Protein; 1140 AA.
XX
AC AAM93971;
DT 30-JUN-1999 (first entry)
XX
DE C. glutamicum pyruvate carboxylase protein.
XX
KW Pyruvate carboxylase; amino acid production; lysine production;
KW threonine production; homoserine production; glutamate production;
KW arginine production; feed additive; condiment; pharmaceutical;
KW fine chemical; ss.
XX
OS Corynebacterium glutamicum.
XX
PN DE19831609-A1.
PD 15-APR-1999.
XX
PF 14-JUL-1998; 98DE-1031609.
XX
PR 04-OCT-1997; 97DE-1043894.
XX
PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI Elkmanns B, Peters-Wendisch P, Sahm H;
XX
DR WP1: 1999-245521/21.
DR N-PSDB; AAX24102.

PT Increasing microbial production of specific amino acids by
PT Increasing activity or expression of pyruvate carboxylase
XX
PS Claim 15; Page 11-15; 18pp; German.
XX
CC This invention describes the isolation of a pyruvate carboxylase
CC from Corynebacterium glutamicum which is used in a novel method for
CC production of lysine, threonine, homoserine, glutamate and/or arginine,
CC variously useful as feed additives, condiments, pharmaceuticals and
CC intermediates for fine chemicals. Increasing pyruvate carboxylase
CC activity increases the yield of microbial production of amino acids
CC of the aspartate and/or glutamate families, e.g. about 50% more lysine,
CC 40% more threonine and 150% more homoserine are secreted into the
CC culture medium.
XX
SQ Sequence 1140 AA;
Query Match 99.9%; Score 5782; DB 20; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSTRSSSTLPAPKKTLLVNRGFIANRAFRALFETAAIVAIYPRDRSGSFHRSFASSEAYR 60
|||||
DB 1 MSTRSSSTLPAPKKTLLVNRGFIANRAFRALFETAAIVAIYPRDRSGSFHRSFASSEAYR 60
QY 61 IGTEGSPYKAYLDIDEIIGAARKVYKADAIYGYGGLSENQAOLARCAENGITFICPTPEV 120
|||||
DB 61 IGTEGSPYKAYLDIDEIIGAARKVYKADAIYGYGGLSENQAOLARCAENGITFICPTPEV 120
QY 121 LDLTGDKSRAVYAAKAGLPVLAESTPSKNIDEIYKSAEGQYPIFYKAVAGGGGRMF 180
|||||
DB 121 LDLTGDKSRAVYAAKAGLPVLAESTPSKNIDEIYKSAEGQYPIFYKAVAGGGGRMF 180
QY 181 VASPELKRKLTEASREAAFGDGVYVERAVINPOIEVOIIGDHGVEVHLYTERCS 240
|||||
DB 181 VASPELKRKLTEASREAAFGDGVYVERAVINPOIEVOIIGDHGVEVHLYTERCS 240
QY 241 LQRRHQKVEIAPAOHLDELRLICADAVKFCRSIGYGAGTVEFLVDEKGNHVFEMN 300
|||||
DB 241 LQRRHQKVEIAPAOHLDELRLICADAVKFCRSIGYGAGTVEFLVDEKGNHVFEMN 300
QY 301 PRIOVEHTVEEYVTVKAOMLAAGATLKEGLTQDKTKTGHAAIQCRITTEDPENG 360
|||||
DB 301 PRIOVEHTVEEYVTVKAOMLAAGATLKEGLTQDKTKTGHAAIQCRITTEDPENG 360
QY 361 FRPDGTITTAAYRSPGAGVRIDGAAGLGEITAHFDSMLVMTGRGSPFETAVARAORAL 420
|||||
DB 361 FRPDGTITTAAYRSPGAGVRIDGAAGLGEITAHFDSMLVMTGRGSPFETAVARAORAL 420
QY 421 AEFTVSGVATNIGFLRALRREDEFTSKRIATGFIADHPHLLQADPADDEGRILLDYADV 480
|||||
DB 421 AEFTVSGVATNIGFLRALRREDEFTSKRIATGFIADHPHLLQADPADDEGRILLDYADV 480
QY 481 TVNKPVGVRPDDVAAPIDKLPINIKDPLPRGSRRLKOLGPAARLDREDDALAVDDTT 540
|||||
DB 481 TVNKPVGVRPDDVAAPIDKLPINIKDPLPRGSRRLKOLGPAARLDREDDALAVDDTT 540
QY 541 FRDAHQSLATRVRSFALKPAAEAVALKLPBELLSEVAMGATYDVAMRFEDPWRDLDE 600
|||||
DB 541 FRDAHQSLATRVRSFALKPAAEAVALKLPBELLSEVAMGATYDVAMRFEDPWRDLDE 600
QY 601 LREAMPNVNIOMLLGRNTVGYTPYDSCRAFYKAEASSGVDFIRIFDALNDVSQMRPA 660
|||||
DB 601 LREAMPNVNIOMLLGRNTVGYTPYDSCRAFYKAEASSGVDFIRIFDALNDVSQMRPA 660
QY 661 IDAVLENTTAAVAEAMAYSGDLSDPNEKLYTDYLLKMAEETVKSAGHILAIKDMAGLLR 720
|||||
DB 661 IDAVLENTTAAVAEAMAYSGDLSDPNEKLYTDYLLKMAEETVKSAGHILAIKDMAGLLR 720
QY 721 PAAVTKLVLTALRREDDLPVHVHTHTDTAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780
|||||
DB 721 PAAVTKLVLTALRREDDLPVHVHTHTDTAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780

QY 781 LSAIVAAFAHTRRDGTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPGRYRHEIPGQL 840
DB 781 LSAIVAAFAHTRRDGTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPGRYRHEIPGQL 840
QY 841 SNLRQAQTALGLADREFLIEDNTAAVNMIGRPVKTPSSKVVGDALHLVAGVDPADF 900
DB 841 SNLRQAQTALGLADREFLIEDNTAAVNMIGRPVKTPSSKVVGDALHLVAGVDPADF 900
QY 901 AADPOKYDIPDSVIAFLRGELGNPGGWPEPLRTALREGSEKCAPLVEPEEOALDA 960
DB 901 AADPOKYDIPDSVIAFLRGELGNPGGWPEPLRTALREGSEKCAPLVEPEEOALDA 960
QY 961 DSKERRNSLNRLFPKPTPEEFLHRRRFGNTSALDREFFYGLVEGRETLIRLPDVRTP 1020
DB 961 DSKERRNSLNRLFPKPTPEEFLHRRRFGNTSALDREFFYGLVEGRETLIRLPDVRTP 1020
QY 1021 LLYRLDAISEPDDKGMRRVAVANNGOIRPKRVDRSVESYATAEKADSSNKGHVAAPEA 1080
DB 1021 LLYRLDAISEPDDKGMRRVAVANNGOIRPKRVDRSVESYATAEKADSSNKGHVAAPEA 1080
QY 1081 GVVTYVVAEGDEVKAGDAVAIIEAMKMEATITASVQKIDRVVVPATKVEGGDLIVVVS 1140
DB 1081 GVVTYVVAEGDEVKAGDAVAIIEAMKMEATITASVQKIDRVVVPATKVEGGDLIVVVS 1140

RESULT 7
AAG93249 standard; Protein: 1140 AA.
AAG93249:
26-SEP-2001 (first entry)
C glutamicum protein fragment mutant P4585.
Coryneform bacterium; amino acid synthesis: vitamin; saccharide;
organic acid synthesis; mutant; muten.
Corynebacterium glutamicum.
Synthetic.
Key Location/Qualifiers
FT MISC-difference 458
FT /note= "wild-type Pro substituted by Ser"
PN EP1108790-A2.
20-JUN-2001.
18-DEC-2000; 2000EP-0127688.
16-DEC-1999; 99JP-0377484.
07-APR-2000; 2000JP-0159162.
03-AUG-2000; 2000JP-0280988.
(KYOW) KYOWA HAKKO KOGYO KK.
Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
Tateishi N, Senoh A, Ikeda M, Ozaki A;
WPI; 2001-376931/40.
Novel polynucleotides derived from Coryneform bacteria, for identifying
mutation point of a gene, measuring expression of a gene, analysing
expression profile or pattern of a gene and identifying homologous gene
Claim 43; Page -: 246pp + Sequence Listing: English.
The present invention provides a number of nucleotide and protein
sequences from the Coryneform bacterium Corynebacterium glutamicum. These
are useful for identifying the mutation point of a gene derived from a
mutant of coryneform bacterium, measuring expression amount and

CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a mutant protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX Sequence 1140 AA:
SQ
Query Match 99.9%; Score 5780; DB 22; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSHTSSTLPFAKKILVANGELAVRAFALETGAATVAIYPRDRGSFHRSAFAVR 60
DB 1 MSHTSSTLPFAKKILVANGELAVRAFALETGAATVAIYPRDRGSFHRSAFAVR 60
QY 61 IGTGSPVKAVIDIDELITGAKKKADAIYPGVEFSENNQLAKECAENGITFGPTPEV 120
DB 61 IGTGSPVKAVIDIDELITGAKKKADAIYPGVEFSENNQLAKECAENGITFGPTPEV 120
QY 121 LDLTGDKSRVATTAARKAGLPLVLAESTPSKNIDEIVKSAEGQYPIFYKAVAGGGGRGMRF 180
DB 121 LDLTGDKSRVATTAARKAGLPLVLAESTPSKNIDEIVKSAEGQYPIFYKAVAGGGGRGMRF 180
QY 181 VASPELRLKATEASREAEAFGDGAVYVERAVINPOHIEVOJIGDHTGEVHLYERDCS 240
DB 181 VASPELRLKATEASREAEAFGDGAVYVERAVINPOHIEVOJIGDHTGEVHLYERDCS 240
QY 241 LQRRHOKVVEIAPQHIDPELROKICADANKFCGSIYOGAGYEFVDEKGNVETEMN 300
DB 241 LQRRHOKVVEIAPQHIDPELROKICADANKFCGSIYOGAGYEFVDEKGNVETEMN 300
QY 301 PRIOVEHTVEEYEVVDLVKAOMLAGATLKEGLTQDKIKTHGALQCRITTEDPNNG 360
DB 301 PRIOVEHTVEEYEVVDLVKAOMLAGATLKEGLTQDKIKTHGALQCRITTEDPNNG 360
QY 361 FRPDGTITTAVRSPGAGVRLDGAOLGSETTAHFDMSLVKMTCRGSDPETAVARAORAL 420
DB 361 FRPDGTITTAVRSPGAGVRLDGAOLGSETTAHFDMSLVKMTCRGSDPETAVARAORAL 420
QY 421 AEFVSGVAATNIGRLRLREEDFTSKRIATGFIADHSHLQAPPADDEGRITLDYADV 480
DB 421 AEFVSGVAATNIGRLRLREEDFTSKRIATGFIADHSHLQAPPADDEGRITLDYADV 480
QY 481 TVNKPGRVPRKDVAPIDKLPNIKDLPLPGSRDRLQOLGPAFARDLREODALAVDTT 540
DB 481 TVNKPGRVPRKDVAPIDKLPNIKDLPLPGSRDRLQOLGPAFARDLREODALAVDTT 540
QY 541 FRDAHOSLATTRVRSFALKPAEAVAKLTPELISVEAMGATYDVAMRFLPEDPWRLDE 600
DB 541 FRDAHOSLATTRVRSFALKPAEAVAKLTPELISVEAMGATYDVAMRFLPEDPWRLDE 600
QY 601 LREAMPVNTQMLRGENTVGYPPYDPSVCARAFKAEASGVDFFRFDALNDVDSQMRPA 660
DB 601 LREAMPVNTQMLRGENTVGYPPYDPSVCARAFKAEASGVDFFRFDALNDVDSQMRPA 660
QY 661 IDAVLENTTAVAEVAMVSGDLSDPNEKLYTLIDVYLLMAEEIVSGAHIIAIKMDALLR 720
DB 661 IDAVLENTTAVAEVAMVSGDLSDPNEKLYTLIDVYLLMAEEIVSGAHIIAIKMDALLR 720
QY 721 PAAVTAKLVTLARREEDLPVHVHTHDATAGGOLATYFAAQAQADAVDASAPLSGTSQPS 780
DB 721 PAAVTAKLVTLARREEDLPVHVHTHDATAGGOLATYFAAQAQADAVDASAPLSGTSQPS 780
QY 781 LSAIVAAFAHTRRDGTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPGRYRHEIPGQL 840
DB 781 LSAIVAAFAHTRRDGTGLSEAVSDLEPYWEAVRGLYLPFESGTPGPGRYRHEIPGQL 840
QY 841 SNLRQAQTALGLADREFLIEDNTAAVNMIGRPVKTPSSKVVGDALHLVAGVDPADF 900


```
Db      |||
841 SMIRAOATATGLADRFELLIEDNYAAVNEMIGRFTVTSKVVGDALATLVAGVDPADF 900
QY      AADPOKYDIPDSVIAFLNGELGNPGGWEPELRTALBGRSGKAPLTFVEPEEQAHIDA 960
Db      |||
901 AADPOKYDIPDSVIAFLNGELGNPGGWEPELRTALBGRSGKAPLTFVEPEEQAHIDA 960
QY      961 DDKERRNSLNLRLPEKPTPEELHRRRRRGNTSALDDREFEYGLVEGRETLLRLPDVTRP 1020
Db      |||
961 DDKERRNSLNLRLPEKPTPEELHRRRRRGNTSALDDREFEYGLVEGRETLLRLPDVTRP 1020
QY      1021 LLVRLDAISEPDDKGMNVANVANOIRPMRVRDSVSSTVATAEKADSSNKGHYAAPPA 1080
Db      |||
1021 LLVRLDAISEPDDKGMNVANVANOIRPMRVRDSVSSTVATAEKADSSNKGHYAAPPA 1080
QY      1081 GVVTVTAAGDEVKAGDAVAITEAKMKMETITASVDGKIDRVVPAATVPEGGDLIVVVS 1140
Db      |||
1081 GVVTVTAAGDEVKAGDAVAITEAKMKMETITASVDGKIDRVVPAATVPEGGDLIVVVS 1140

RESULT 8
AAU98052
ID      AAU98052 standard; Protein; 1140 AA.
AC      AAU98052;
DE      27-AUG-2002 (first entry)
XX      Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
XX      Feedback-resistant; pyruvate carboxylase; enzyme;
XX      aspartic acid feedback inhibition resistant; mutant; muten.
OS      Corynebacterium glutamicum.
XX      Synthetic.
FH      Key
FH      Location/Qualifiers
FT      Misc-difference 1 /note= "Wild-type Met substituted by Val"
FT      Misc-difference 153 /note= "Wild-type Glu substituted by Asp"
FT      Misc-difference 182 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 206 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 227 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 455 /note= "Wild-type His substituted by Arg"
FT      Region /note= "Wild-type Ala substituted by Gly"
FT      Misc-difference 1116 /note= "Specifically claimed in claim 18"
FT      Misc-difference /note= "Wild-type Asp substituted by Glu"
XX      WO200231158-A2.
XX      18-APR-2002.
XX      12-OCT-2001; 2001WO-US31893.
XX      13-OCT-2000; 2000US-239913P.
XX      (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX      Hanke PD;
XX      WPI: 2002-463267/49.
XX      Novel mutated, feedback resistant pyruvate carboxylase enzyme
XX      polypeptide, useful for producing amino acids e.g. L-lysine,
XX      L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
XX      L-isoleucine
```

```
PS      Claim 1: Page -: 42pp; English.
XX      The present invention relates to a new mutated, feedback-resistant
CC      pyruvate carboxylase enzyme. The invention is useful for producing an
CC      amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC      by culturing a host cell in a suitable media and separating the amino
CC      acid from the medium. The vector of the invention is useful for
CC      replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC      resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC      replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC      a selectable marker gene through homologous recombination to form a first
CC      recombinant strain, and replacing the selectable marker gene in the
CC      first recombinant strain, with feedback resistant pyruvate carboxylase
CC      gene through homologous recombination to form a second recombinant
CC      strain, where the homologous recombination in the above steps, occurs
CC      between the host cell and the vector. The feedback-resistant pyruvate
CC      carboxylase enzyme is resistant to feedback inhibition from aspartic
CC      acid. The present amino acid sequence represents the mutant
CC      feedback-resistant pyruvate carboxylase enzyme of the invention.
CC      Note: The present sequence is not shown in the specification but is
CC      derived from the wild-type feedback-resistant pyruvate carboxylase
CC      enzyme (AAU98053) given in figure 2 of the specification.
XX      Sequence 1140 AA;
```

```
Query Match          99.5%; Score 5759; DB 23; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 MSTRSSITLPFAFKKLLVNRGEIYAVRAALETCAAVAIYPRDRSGSFHRSFSEAVR 60
Db      1 VSTRSSITLPFAFKKLLVNRGEIYAVRAALETCAAVAIYPRDRSGSFHRSFSEAVR 60
QY      61 IGTEGSPYKAYLDIDEITGAARKKVAADIVYGGYGLSNAOALARCANGTFTGPTPEV 120
Db      61 IGTEGSPYKAYLDIDEITGAARKKVAADIVYGGYGLSNAOALARCANGTFTGPTPEV 120
QY      121 LDLTGDKSRAYTAARKAGLPVLAESTPSKNIDETVKSAGQTYPIFYKAVAGGGGGRMR 180
Db      121 LDLTGDKSRAYTAARKAGLPVLAESTPSKNIDETVKSAGQTYPIFYKAVAGGGGGRMR 180
QY      181 VASPELKLKLTASREAEAFGCGAVYVERAVINPOHIEVOILLGDRHGEVHHLYERCCS 240
Db      181 VASPELKLKLTASREAEAFGCGAVYVERAVINPOHIEVOILLGDRHGEVHHLYERCCS 240
QY      241 LORRHOQVETAPAOHLDPELRDRICADAVFCRSIGYGAGTVEFLVDEKGNHFTFPMN 300
Db      241 LORRHOQVETAPAOHLDPELRDRICADAVFCRSIGYGAGTVEFLVDEKGNHFTFPMN 300
QY      301 PRIOVEHTVTEEVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
Db      301 PRIOVEHTVTEEVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
QY      361 FRPDGTITAYRSFGAGVRLDGAALQGETTANFDSMLVMTKRGSGFETAVARAQAL 420
Db      361 FRPDGTITAYRSFGAGVRLDGAALQGETTANFDSMLVMTKRGSGFETAVARAQAL 420
QY      421 AEFYSGVATNIGFLRALREDEFTSKRIATGFTADHGHLLQAPPADDEQGRITDYADV 480
Db      421 AEFYSGVATNIGFLRALREDEFTSKRIATGFTADHGHLLQAPPADDEQGRITDYADV 480
QY      481 TVNKPVGVRPDVAPIDKLPNIDPLPRGSSRDLKQLGPAARFRLREODALAVYDTT 540
Db      481 TVNKPVGVRPDVAPIDKLPNIDPLPRGSSRDLKQLGPAARFRLREODALAVYDTT 540
QY      541 FRDAHQSILATRVVSFALKPAEAVAKITPELLSVEAMGATTVYANRFLFEDPDWRIDE 600
Db      541 FRDAHQSILATRVVSFALKPAEAVAKITPELLSVEAMGATTVYANRFLFEDPDWRIDE 600
QY      601 LREAMPNVNIDMLRGRTVGYTPYPSVCAPAFKEAASSGVDFRIFDALNDVSOBRPA 660
Db      601 LREAMPNVNIDMLRGRTVGYTPYPSVCAPAFKEAASSGVDFRIFDALNDVSOBRPA 660
```


Db 558 FRDAHOSLATTRRSALRPAAEVAKLPELLSVEAWGATYDAMRFLFEDPWRDLDE 617
QY 601 LREAMPNVNIOMLLRGNTVGYTPYPDSYCRAFVKEAASSGVDFRIFDALNDVSOBRPA 660
Db 618 LREAMPNVNIOMLLRGNTVGYTPYPDSYCRAFVKEAASSGVDFRIFDALNDVSOBRPA 677
QY 661 IDAVLETNTAVAEVAAAYSGDLSDPNEKLYTLTDYLYKMAEEYKSGAHLIAIKDMAGLLR 720
Db 678 IDAVLETNTAVAEVAAAYSGDLSDPNEKLYTLTDYLYKMAEEYKSGAHLIAIKDMAGLLR 737
QY 721 PAAVTKLYTALRREFLPVHVHTHDFAAGOLATYFFAAAGADAVGASAPLSGTSOPS 780
Db 738 PAAVTKLYTALRREFLPVHVHTHDFAAGOLATYFFAAAGADAVGASAPLSGTSOPS 797
QY 781 LSAIVAAFAHTRRDGTLSLEAVSDLEPYWEAVRGSLPPESGTGPCTGRVYRHEIPGQOL 840
Db 798 LSAIVAAFAHTRRDGTLSLEAVSDLEPYWEAVRGSLPPESGTGPCTGRVYRHEIPGQOL 857
QY 841 SNIRAAQATALGLADREPLEEDNYAAVNEMLGRPTKVTSPSSKYVGDIALHLVGAGVDPADP 900
Db 858 SNIRAAQATALGLADREPLEEDNYAAVNEMLGRPTKVTSPSSKYVGDIALHLVGAGVDPADP 917
QY 901 AADPOKYTIDIPDSYIAFLRGELGNPGGWPEPLRTRALBGRSGKAPLITEVPEEQAHDLA 960
Db 918 AADPOKYTIDIPDSYIAFLRGELGNPGGWPEPLRTRALBGRSGKAPLITEVPEEQAHDLA 977
QY 961 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVEGRETLIRLPDVTRTP 1020
Db 978 DSKERRNSLNRLLFPKPTPEEFLEHRRRGNTSALDDREFFYGLVEGRETLIRLPDVTRTP 1037
QY 1021 LTVRLDAISEPDDKGRNVYANVNGOIRPMRYRDSVESVYTAIAEKADSSNKGHVAAPPA 1080
Db 1038 LTVRLDAISEPDDKGRNVYANVNGOIRPMRYRDSVESVYTAIAEKADSSNKGHVAAPPA 1097
QY 1081 GVVTVYVABGDYKAGDAVAIIEAMKMEATITASVDGKIDRVYVPATVVEGGDLIVVVS 1140
Db 1098 GVVTVYVABGDYKAGDAVAIIEAMKMEATITASVDGKIDRVYVPATVVEGGDLIVVVS 1157

RESULT 10
AAB83180
ID AAB83180 standard; Protein; 1139 AA.
XX
AC AAB83180;
XX
09-JUL-2001 (first entry)
XX
DE Corynebacterium thermoaminogenes pc protein.
XX
KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW thermotolerant; acea; accBC; dtsR1; dtsR2; pfk; scrB; gluaBCD;
KW pdha; pc; ppc; acn; 1cd; 1pd; odha.
XX
OS Corynebacterium thermoaminogenes.
XX
PN W0200125447-A1.
XX
PD 12-APR-2001.
XX
PF 04-OCT-2000; 2000WO-JP06913.
XX
PR 04-OCT-1999; 99JP-0282716.
PR 01-NOV-1999; 99JP-0311147.
PR 21-APR-2000; 2000JP-0120687.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, kimura E;
PI Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
PI Sugimoto S;
XX
DR WPI: 2001-300170/31.
DR N-PSDB; AAF87437.

XX
PT Proteins and their DNA useful for microbial production of L-amino acids
PT
PS Claim 9; Page 132-135; 215pp; Japanese.
XX
CC The present sequence is provided in a specification relating to genes
CC encoding thermophilic amino acid biosynthesis system enzymes of
CC the thermotolerant bacterium Corynebacterium thermoaminogenes.
CC The novel proteins retain at least 30% isocitrate lyase activity
CC after heating at 500C for 5 minutes. DNA fragments encoding the
CC enzymes were isolated from a Corynebacterium thermoaminogenes
CC chromosomal DNA plasmid library by PCR. The DNA may be used for
CC developing strains of amino acid producing microorganisms.
XX
SQ Sequence 1139 AA;
Query Match 91.6%; Score 5302.5; DB 22; Length 1139;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 1036; Conservative 43; Mismatches 58; Indels 1; Gaps 1;
QY 3 TTTSSTLPFAFKKILVANKREIVAVRAALEGAATVAIYPREDGSHRSFASAVRIG 62
Db 3 TTTPSTLPFAFKKILVANKREIVAVRAALEGAATVAIYPREDGSHRSFASAVRIG 62
QY 63 TEGSPVKAVIDIDEITIGAAKKVKAADIYPGYGFLSENNOLARECAENGITFGPTPEYVD 122
Db 63 TEGSPVKAVIDIDEITIGAAKKVKAADIYPGYGFLSENNOLARECAENGITFGPTPEYVD 122
QY 123 LTGDKSRAYTAARKKAGLPVLAESTPDSKINIDEIVKSAEGQYPIFYKAVAGGGGRMPVA 182
Db 123 LTGDKSKAVSAKKAGLPVLAESTPDSKINIDEIVKSAEGQYPIFYKAVAGGGGRMPVE 182
QY 183 SPDELRLKATESSAREAAFGGAYVERAVYNPNHIVQILGHTGEVNHLYEEDCSLQ 242
Db 183 KREDIRELAREASREAEAFGGSVYVERAVYKPPHIEVQILGHTGTVIHLIYERDCSLQ 242
QY 243 RRHOKVVEIAPAOHLDPELRDRICADAVKFCISIEYOGAGVVEFLVDKGNHVFEMPR 302
Db 243 RRHOKVVEIAPAOHLDPELRDRICADAVKFCISIEYOGAGVVEFLVDAGNHVFEMPR 302
QY 303 IOVEHTVTEEYTVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFR 362
Db 303 IOVEHTVTEEYTVSVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNSNFR 362
QY 363 PDGTGTTAVRSPGCGAVRLDGAOLGGEITAHFDSMLVKMKCRGSDFTAVSRAQORALAE 422
Db 363 PDGTGTTAVRSPGCGAVRLDGAOLGGEITAHFDSMLVKMKCRGSDFTAVSRAQORALAE 422
QY 423 FTVSGVATNIGFLRALRREDFTSKRIATGFIADPHLLQAPPADDEGRIIDYLAADV 482
Db 423 FNVSGVATNIGFLRALRREDFTSKRIATGFIADPHLLQAPPADDEGRIIDYLAADV 482
QY 483 NKPHGVRKDYAAPIDKLPNTKDLPLPGSGDRKLKOLGPAAPFANDLREQDALVDTYFR 542
Db 483 NKPHGERP-ETARPLETEKLEVENIPLPGSHDRKLKOLGPEGFARDLREQDALVDTYFR 541
QY 543 DAHOSLATTRRSFALRPAAEVAKLPELLSVEAWGATYDAMRFLFEDPWRDLDELR 602
Db 543 DAHOSLATTRRSFALRPAAEVAKLPELLSVEAWGATYDAMRFLFEDPWRDLDELR 601
QY 603 EAMPNVNIOMLLRGNTVGYTPYPDSYCRAFVKEAASSGVDFRIFDALNDVSOBRPAID 662
Db 602 EAMPNVNIOMLLRGNTVGYTPYPDSYCRAFVKEAASSGVDFRIFDALNDVSOBRPAID 661
QY 663 AVLETNTAVAEVAAAYSGDLSDPNEKLYTLTDYLYKMAEEYKSGAHLIAIKDMAGLLRPA 722
Db 662 AVLETGTSVAEVAAAYSGDLSNPGEKLYTLTDYLYKMAEEYKSGAHLIAIKDMAGLLRRA 721
QY 723 AVTKLYTALRREFLPVHVHTHDFAAGOLATYFFAAAGADAVGASAPLSGTSOPS 782
Db 722 AAKLYTALRREFLPVHVHTHDFAAGOLATYFFAAAGADAVGASAPLSGTSOPS 781

|||||
Db 378 FRPDGTITAYRSPGAGVRLDGAAGGEBITAHFDSMLVKMTCRGSDPETAARARAL 437
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADV 480
Db 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADV 497
QY 481 TVNKPFGVPRKDVAAPIDKLPIKIDLPPLRGSRDR 515
Db 498 TVNKPFGVPRKDVAAPIDKLPIKIDLPPLRGSRDR 532
RESULT 12
AAU00511
ID AAU00511 standard; Protein; 1148 AA.
XX AAU00511;
XX
XX 07-SEP-2001 (first entry)
XX
DE Bacillus subtilis pyruvate carboxylase enzyme A.
XX
XX
XX Pyruvate carboxylase A; pycA; oxaloacetate; Escherichia;
KM L-amino acid production; fermentation.
XX
OS Bacillus subtilis strain 168.
XX
PN EPI092776-A1.
PD 18-APR-2001.
XX
PE 05-OCT-2000; 2000EP-0121763.
XX
PR 14-OCT-1999; 99RU-0121636.
XX
XX (AJIN) AJINOMOTO CO INC.
PA Gustylier MM, Kozlov VI, Piltisyn LR, Altman IB, Voroshilova EB;
PI Iomantas YAV, Yampolskaya TA;
PI Iomantas YAV, Yampolskaya TA;
DR WPI; 2001-309819/73.
DR N-PSDB; AAS01509.
XX
XX New bacterium from the genus Escherichia containing a gene encoding for
PT pyruvate carboxylase useful for producing higher concentrations of
PT L-amino acids
XX
XX Claim 4; Page 17-21; 28pp; English.
XX
XX The present sequence represents Bacillus subtilis pyruvate
CC carboxylase enzyme A. The pycA gene encodes for pyruvate carboxylase
CC which catalyses the carboxylation of pyruvate to form oxaloacetate.
CC Transformation of the Bacillus subtilis pycA gene into a bacterium
CC belonging to the genus Escherichia results in the bacterium showing
CC L-amino acid productivity. The invention provides a novel method for
CC producing an L-amino acid by fermentation. The method involves
CC culturing the bacterium in a medium and producing and accumulating
CC the L-amino acid in the medium. The new bacterium harbouring the gene
CC coding for an enzyme having pyruvate carboxylase activity is useful
CC for producing higher concentrations of L-amino acids in vitro than
CC prior art.
XX
XX
SQ Sequence 1148 AA;
Query Match 43.6%; Score 2524; DB 22; Length 1148;
Best Local Similarity 46.0%; Pred. No. 1.1e-176;
Matches 533; Conservative 186; Mismatches 392; Indels 48; Gaps 13;
QY 11 AFKKIIVANRGELAVRAPRALETAATATATYPRREDRGSFHSPASEAVRITGESSPYKA 70
Db 5 SIQKLVAVANRGELAIRIRFRACTELNIRITAVAVSKEDSSGSHRYKKADEAVLYVEGKRPPIDA 64
QY 71 YLDIDELIGAAKKVAKADAIYPGYGLSENAQIARCAENGITGIGPTPEVLDLTGDKSRA 130

|||||
Db 65 YLDIEGIIIDIAKKNKVDALHPGVGLSENHFAFRCBEEBIVTIGKSEHLDHFGKQVKA 124
QY 131 VTAAKKAGLPVLAEST-PSKNIDEIYKSAEGQITPIYKAVAGCGRGMRFAVSPDLRK 189
Db 125 REQAEKAGIPVIGSGPAPETLEAVBFGGANGYPIIKIASLGGGGRGRKIRYSESEVKE 184
QY 190 LATEASREAAFGADGAVYVERAVINPQHEVOIILDHNGEYVHLTERPDSIORRQKTV 249
Db 185 AYERASSEAKAAGNDEVIYEKLEIENPKHIEVOYIDKQGNVYHLERDCSVORRQKTV 244
QY 250 ETAPAOHLDELDRICADAVKFCRSIGVAGTVEFLVDKGNHVFIEIMPNIQVEHTV 309
Db 245 EVAPSVLSPELBDQICEAAVALAKWNYINAGTVEFLV-ANNEFFYIEVNPVQVEHTI 303
QY 310 TEEVTEVDLYKQKMRILAAGATL--KELGLQDK-ITKGAALOCRTTDEDPNNGFPDNG 366
Db 304 TEMITGVVDIVQVQILVAQGHSLHKKVNIPEODIPTIIGYAIQSRYTDEDPODFMPDVG 363
QY 367 TITAYRSPGAGVRLD-GAAQLGGEITTAHFDSMLVKMTCRGSDPETAARARALAEFTV 425
Db 364 KIMAYRSGGFGVRLDTGNSFOGAVITPYIDSLVSLWTALTFEQAAMKMRNLOEFTI 423
QY 426 SGVATNIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADVTVN-- 483
Db 424 RGIKTNIPLENVAKHEKFLTGQYDTSFIDTTPLEFNPKOKDRGKMLTIGNYTVNGF 483
QY 484 -----KPHGVPRKDVAAPIDKLPIKIDLPPLRGSRDLKQIGPAAPRDLREDDALAV 536
Db 484 PGIGKKKEKPEKPLGVKVDVDOQ-----ARGKQILDEKGAEGLANMWEKSVLL 536
QY 537 TDTTERDAQSLATFVRSEFALKPAEAVAKITPELTSVEANGATYDVAAMFLEPEDPD 596
Db 537 TDTTERDAQSLATFVRSEFALKPAEAVAKITPELTSVEANGATYDVAAMFLEPEDPK 596
QY 597 RIDELEAMPVNNIOMLGRNTVGTTPYPDSCRAFEVKAASSGVDFIRFDALNDVSQ 656
Db 597 RLEDLKEVNPTEFLMRLSSNAVGYTNPDVNIKEFYKQASQSGIDVFRIDPSLWVG 656
QY 657 MRPAIDAVLETMTVAEVMAYSGDLSNENKLYTLIDYIKMAEFYVSGAIIILIKMA 716
Db 657 MTLAIDAVDVG-KVAEAAICTGDLIDNRKRYDILATYTSMAKLEDAAGAILILIKMA 715
QY 717 GLLRPAAYTKLTALREFDLFVHVHTHDYAGQIATYFAAQAQADAVDGAASAPLGGTT 776
Db 716 GLIKPOAAVEIYSAIKETIDIPVHLHTDTSNGIYMAKAVEAGVDIIDVSSMAGLT 775
QY 777 SOPSLSAIYAAFAHTRROTGLSLEAVSDLEPTEWEAVRGLIYLFESGTFEPTGRVYRHET 836
Db 776 SOPSASGEYHAMEGNDREPMNVQVELLSQYVESVRYKYSFEESGMKSPHTEIYEHMP 835
QY 837 GQOLSNLRQAATALGLADREFELIEDNYAAVNMELRPTKVPSSRYVGDIALHLYGAGVD 896
Db 836 GGQYENLQQAAGVGLGRBMNVKEMRYRVNMFEDYIKVTPSSRYVGDIALHLYQNNLT 895
QY 897 PADFPADPOKIDIPDSYAFELRGELGNPGGMPPELRTALRGSEKAPLVEPEE-- 953
Db 896 EKDYVEKESDLPFSPVVELEFGNIGOPHGGEPEKILQILIKGQF-----PIYVRGPELE 951
QY 954 -----EQAHLADDSKERRNSLNLPLPKPEEFLNRRRRGNLSALDDREF 1001
Db 952 PVSFEALIOERKEQNNLEISD-----QDAVAALYKVFYTDYKTTESYGDLSLDTPTFF 1007
QY 1002 YGLVEGRETLIRLPDVRTPLVLRILDAISEPDDKGRNRYVANVNGQIRPMRYDRSVESVT 1061
Db 1008 YGMTLGEIEIYEIERGKI-LIVKLISIEBPQDARRVYFELNCGPRRVYIKDESIKSSV 1066
QY 1062 ATAERADSSNKGHVAAPRAGVYTVVAA-GDEVKAGDAVAILIEMKMEATTATYADGKID 1120
Db 1067 QERLKADRTNPSHIAASMPGTVIKVLAEGTKVNGDHLINEMKMETVQAFPSGTIK 1126
QY 1121 RVVVPAAATKVEGGDLIVVY 1139

Db 1127 QVHVKNGEPIOTGDLLEI 1145

RESULT 13

AB47612 standard; Protein: 1146 AA.

XX AB47612:

AC AB47612:

XX 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #316.

XX

KM Antibacterial: gene therapy: vaccine: biosynthesis; biodegradation:

KM vitamin B12; bacterial infection; disease.

XX

OS Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCt-2001.

XX 11-APR-2001: 2001WO-FR01118.

XX 11-APR-2000: 2000FR-0004629.

XX (INSP) INST PASTEUR.

PA Buchteser C, Frangeul L, Couve E, Rusnok C, Fsihi H, Deloux P;

PI Dussauget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JB;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Bommert E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Meduendo E, De Padlos B, Wehlund J, Kaerst U, Entlian K, Hant J;

PI Rose M, Voss H;

PI

DR WPI; 2002-010914/01.

XX

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of Listeria and related bacterial infections, and

PT related polypeptides -

PS Claim 6: SEQ ID No 317; 192pp: French.

XX

CC The present invention relates to the genome sequence of Listeria

CC monocytogenes EGD-e (see AB403041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in L.

CC monocytogenes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate L. monocytogenes-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by L.

CC monocytogenes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WPIPO

CC at ftp.wipo.int/pub/published_sequences.

XX

SQ Sequence 1146 AA:

Query Match 43.1%; Score 2493; DB 23; Length 1146;

Best Local Similarity 45.8%; Pred. No. 2,1e-174;

Matches 526; Conservative 189; Mismatches 408; Indels 30; Gaps 11;

OY 13 KKLIVANRGEIIVRAFALEGTGAIVAIIPREDRGSFHRSPASEAVRIGTSGSPKAVYL 72

DB 5 KKLIVANRGEIIVRAFALEGTGAIVAIIPREDRGSFHRSPASEAVRIGTSGSPKAVYL 64

OY 73 DIDEIIGAARKKADAIYPGYGLSENALARECAENGITTEIPPEVLDTGDKRAVT 132

DB 65 DIENITETAKESGADALHPGIGFLESENIEFARCEOGITTFVGKSHLDLFGDKITAKE 124

OY 133 AAKRAGLPVLAEST-PSKNIDEIVKSAEGQTPYFVAVAGGGGRGFRVAPSDRLKLA 191

DB 125 QALLADIPVIGSGNPVAGIKVEFEPEKKNYPLMIKASLGSGGRGRVRESKREHVESF 184

OY 192 TEASREAAFGGAGVYVERAVINPOHIEVOILDTGTEVHLTERDCSLORRQKVEI 251

DB 185 ERASSEAKAAFGNDEYVERKCVNMPKHIEVOILDTGTEVHLTERDCSLORRQKVEI 244

OY 252 APACHLDELDRICADAVKRCRISIGOGATVEFLVDEKGNHIFIMNPIOVHEHYTE 311

DB 245 APCNAITSELNRIICDAAVKLMKNVDYINGTEVEFL-EGDDFYFIEVNPVOVEHYTE 303

OY 312 EYFEVDLVKAKOMRLAAGATLKEGLT---ODKITHGAALQCRITTEDPNNGFRPDGTI 368

DB 304 MITGIDIVGSLFIADGVALHDOLVAIPKQEDIHHSAGISRTTEDPLNFMPPDGRV 363

OY 369 TAYRSPGAGVRD-GAOLGGEITAHFDSKLVKMTGRGSDFEYAVARAQALAEFTVSG 427

DB 364 DYRSTGFGVRLDAGNGFQGVVTPPYDSLVLVLCYMGMTFEQATKRMRLTEFRIG 423

OY 428 VATNIGFLRLAREDEFTSKRIATGFIADHPHLLQAPPADEQRIIDYADVTVNKPFG 487

DB 424 VKTNIPLFLNVRHPDASGNVNTSFDTTTELEKFIHRIKGTITLIGNVNKPFG 483

OY 488 VRPDPVAPIDIKLPINIKLP---LPRGSRDLKQLGPAAFARDLREDAVALVDTTFRD 543

DB 484 IKHND-KPYAERFLKPIRGISQISPTKQIOLDAKPEGVDVAKQOEVLDTDLTRD 541

OY 544 AHOSLATRVRSKXIFOVADAMAHLLPMSFEMKAGATEDVAYRFLNEPVRVLETLRK 601

DB 542 AHOSLATRVRSKXIFOVADAMAHLLPMSFEMKAGATEDVAYRFLNEPVRVLETLRK 601

OY 604 AMPNVNIOMLIRGNRTGYTPYPSVCRAPKFAASGVDFEFLDLNDVSONRPAIDA 663

DB 602 QIPNVFOMLIRGNNAAGYKPNVIREFKQASGVDFRFDVSNMKGVEVSDA 661

OY 664 VLENTVAEVAAMVSGDLPNEKLYTLDYLLMAEIVKSAHILAIKDMAGLLRPA 723

DB 662 VREAG-KYVEALIGYTDIDDDTRKTYTIDYKMAKEIVAGNHILIGKMAAGLLRPA 720

OY 724 VTKLVTLRREEDLPVHVHTHDYAGOLATYFAAQAQAGADVAGASAPLSGTSQPSLSA 783

DB 721 AYRLIGELKIDVDVPIHLHTHDSGNGIYTYAAAVSAGVDIVDASSAMSGATQSPMTG 780

OY 784 IVAFARTRDPTGSLSEAVSDLEPYMEVRLYLPEFSGPTGPRGRYRHEITPGGOLSNL 843

DB 781 LVYGLVGNRQTNLDAQNSQIINHMYEDVRYHYKDFPNALNSPQTEVYIHEMPGQYTNL 840

OY 844 RAQATAGLADREFLEIDNVAAVNEMIGRPKVPSSKVVGDLLHLVAGAGVADPADPAD 903

DB 841 QQQAIVAGLGRMVEKMYTVVQMGDVIKVPSSKVVGDLLHLVQVQNLSEDEYK 900

OY 904 POKIDIDPSYAFIRGELGNPGRGMPERLTRALEGSEKAPLITEVPEEQAHLDADS 963

DB 901 GDTIDPDSVIEFPMGEIGQPYGGFPEKLOLVYKGR---PLTDRGALMEVNPVDV 956

OY 964 K-----ERRNSLNRLFPKPTPEEFLEHRRRGNALSALDREFYGLVGRGTLI 1012

DB 957 KAELEKMGVEPTKDVISTILYKVLVDYODMINKGDTVLTDPFFYKMRIGETIEV 1016

OY 1013 RLDPVTRPLVRLDAISEPDDKGRNVAVANGQIRPMVRDRSVESYATATKADSSNK 1072

DB 1017 ELEKGL-LIKLNSISEPADTRGVYFELNQGPREINIQDMVQSGVYARRKIDPTNP 1075

OY 1073 GHVAAPFAG-VVYTVVAGDEEVKAGDAVAITEAKMEATITASVDGKIDRVVPAATKE 1131

DB 1076 EHVAGATMGSVIQVYVVKGDSYKGPDLITLAKMKMETTIOAPDEGVSSIYVSDGPTIE 1135

Db 1060 NWHTNANKPRKADSNPSHIGAMGVSYTEVKSVCETVKANOPLLITEAMMETTIOAP 1119
 QY 1115 VDGKIDRVVPAATKVEGDLIVVY 1139
 Db 1120 FDGVAKOVTYNNGTIATGDLITET 1144

RESULT 15
 AAU35213
 ID AAU35213 standard; Protein: 1142 AA.

XX AAU35213;
 AC 13-FEB-2002 (first entry)
 DT
 XX Enterococcus faecalis cellular proliferation protein #500.
 DE
 XX Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX Enterococcus faecalis.
 OS
 XX WO200170955-A2.
 FN
 XX 27-SEP-2001.
 PD
 XX 21-MAR-2001; 2001WO-US09180.
 PE
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 PR
 XX (ELIT-) ELITRA PHARM INC.
 PA
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 N-PSDB: AAS53072.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 10806; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1142 AA;
 SO

Query Match 42.7%; Score 2472.5; DB 22; Length 1142;

Best Local Similarity 45.8%; Pred. No. 6,8e-173;
 Matches 532; Conservative 185; Mismatches 386; Indels 59; Gaps 17;

QY 13 KLIIVANGETIAVAFRALETGAATVAIYREDGSTRHSASAVNIIGEGSPVAVL 72
 Db 2 KVLIVANGELAIRIFRACTELDIRTAIYAEDYSVHRFADADAYLVGKEKKIEAYL 61
 QY 73 DIDEIIGAIAKKVADAIYGYGFLSENOQLARECAENGTITGPTPEVLDLTGDSRAVY 132
 Db 62 DIENITQIAKKSAGDAIHFGYGFLESENLRFAECCEEGIIYVGPETHLIDFGDKYAKE 121
 QY 122 AAVAAGIASIPGSDGPVATVEEVAVFGEHGPIMIKALGGCGRMVADAKARGY 181
 Db 133 AAKKALPVLAEET-PSKNIDEIVKSAEGQYPIFKAAGGCGRMRFVASPDELRL 191
 QY 122 AAVAAGIASIPGSDGPVATVEEVAVFGEHGPIMIKALGGCGRMVADAKARGY 181
 Db 192 TEASREAEAGDCAVYVERAVINQIHIEVOILGHTGEVHLYRDCSLDRHKKVVEI 251
 QY 182 ERAKSEAKAAGSDSEVVEYKYSNPKHIEVOILGDHGNVHLFERDCSVORRHKVVEY 241
 Db 252 APAOHLDELDRICADAVKFCRSIGYGAGVEFLVDKGNHVFTEKMPRIOVETYTE 311
 QY 242 APCVSMNEQRAITCSAAVQOLMAHVGYNAGTVEFLV-EGDOFYIEVNPQVYVHTITE 300
 Db 312 ETEVDLVKAQMLAAGATL-KELGL-TQDKIKTHGAALOCRTTEDPNNGRPDTCIT 369
 QY 301 MITDIDIYISQIQAQGLDLHKHDMHLPKONELTLKGAIIQCRITTEDPLNQFMPDTGKID 360
 Db 370 AYRSPGAGVRLD-GAOLGCEITTAHFPSMLVKMTKCRSDTEFAVARORALAEFTVSGV 428
 QY 361 TYRSPGCGFVRLDVGNAVSGYAVPYFSLVLYKVCVTHGFSFOAISKMKRCKEIRIGV 420
 Db 429 ATNIGELFALLREDEFTSKRLATGFIADPHLLAPADDDEGRILDLADTVTKPKGV 488
 QY 421 KTNIPFLQNVVSYPAFOSGEAKTFTIDTPELFEEPRMRDGNKMTKTIGETVAVGFGPI 480
 Db 489 R-----PNDVAAPIDKLPINIKLPLRSGSRDLKQLGPAFAFRLDREODALAV 536
 QY 481 ERTEKKYFEAPRPVDIEVP-EKVITANL-----LDAQATVAIVDMKNGESVLM 530
 Db 537 TDTTFRDAHOSLATAVRVSFALKPAAEVAKLT---PELLSVEAMGATTDVAMRFLPE 592
 QY 531 TDTTFRDAHOSLATAVRVTRTDFK---AIAGLTIDVALBELFSEEMGATGDDVAYRFLTE 586
 Db 593 DPMRDLDELREAMPVNTOMLLRGRNTGYTPYPSVCARAFKAASGVDFRFFDLN 652
 QY 587 DPMOHLKRIQDMPHILLQMLEFRGSNAVGYONTYPNVNIEETIKESARCGVDVFRFFDLN 646
 Db 653 DVSGMRPAIDAVLENTAVAEVAMAYSGDLSDPNEKLTLYLTYLKMAEIVKSGAHILAI 712
 QY 647 WIPQEKESIQVVRDNG-KIAEALICYTGDIINDPAKKNVQYILDMAKELEMLGAIITAI 705
 Db 713 KDMAGLPRPAVTKVTLRLREBFLPVVHTHTDAGGOLATYPAANOAGADAVGASAPL 772
 QY 706 KDMAGLPRPAVTKVTLRLREBFLPVVHTHTDAGGOLATYPAANOAGADAVGASAPL 765
 Db 773 SGTTSQPSLSIVAAFAHTRBDTGLSLAVSODLEBYWAVGVLYLPEFSGTPGPGRYR 832
 QY 766 SGATSPSPMSNSIYALVNGERTPTINIDNAOKINHWDVRYIYOPFENGILNAPTEVYM 825
 Db 833 HEIPGQSLSNLRQAOTATGALDRELELDNAAVNMELGRPTKYVPSKVVGDALHLVYG 892
 QY 826 HEMPEGQYSNLOQAQAKAGLGRHMEIKKMTHTVLMHGDLYKVPSSKVVGDALHFMVQ 885
 Db 893 AGVDPAFDAADPQKYIDPDSYIAFLRGELGNPFGGMEPLRLTRALEGSECKAPLVEVE 952
 QY 886 NHLTEQDVPYARGEEISFEESVYTFEGDGLGQVPGFPELRIILIKR---PAFTERPG 941
 Db 953 EEOAHLDDDSKER-----RNSLNLNLPKPTFEELHERR---RGNISALDDR 998
 QY 942 DLAAVDPAKVOEELAEKIGYQPKLEEVLSYLMRP---QVLETKQKETEYGDITLDTLP 998
 Db 999 EEFYGLVGRETLIRLPDVRTPLVRLDAISEPDDKGRNVAVANVNGOIRPMRVADRSEVE 1058
 QY 1115 VDGKIDRVVPAATKVEGDLIVVY 1139
 Db 1120 FDGVAKOVTYNNGTIATGDLITET 1144

Thu Sep 25 08:53:46 2003

us-09-974-973a-19.rag

Page 16

Db 999 TFGNGHQGEHLEUVQIERGKTLIIRDEIDGEPIIDGNRVLFEMUNQOREVLVKDASIK 1057

Qy 1059 SVTATKAKDSNNCHVAAPFAGVYVYVTAEGSEYKAGDAVAIIEMKKREATTASVDG 1117

Db 1058 SAVOVKQKAEPETNEQIGATMSGVLOLVAKRGDVKERGQPLITTEAMKMETTEARFAG 1117

Qy 1118 KIDRVVYPATKVEGGDLIVVY 1139

Db 1118 TVDHIIVEEGEAISSGDLILEV 1139

Search completed: September 24, 2003, 15:55:26
Job time : 68.5263 secs